ACVM - 2. Assignment - Mean-Shift tracking

Nejc Ločičnik, 63180183, nl4952@student.uni-lj.si

I. Introduction

In this assignment, we delve into Mean-Shift tracking. Initially, we examine the mean-shift algorithm and its parameters using simple response images. Subsequently, we integrate the algorithm into a basic tracker and test it on multiple image sequences to enhance our understanding.

II. Experiments

A. Mean-shift parameters

In implementing the mean-shift algorithm, we experiment with different parameters, notably the type of kernel (Epanechnikov or Gaussian) and its size. These parameters significantly influence the algorithm's ability to converge to an optimum. For instance, the Gaussian kernel may converge to a local optimum due to its weighted nature, while the Epanechnikov kernel, with its constant derivative, is less prone to such issues.

Regarding kernel size, it plays a crucial role in determining convergence. A larger kernel size increases the likelihood of proper convergence by ensuring that areas with no gradient are traversed effectively. However, overly large kernels can hinder convergence once the algorithm nears the optimum, introducing irrelevant noise from outer areas. Although reducing the kernel size iteratively could improve accuracy and speed up convergence, the gains are marginal compared to lost computational efficiency (certain parts now need to be inside the iterative loop).

Typically, the algorithm's identified optimum is off by 1-2 pixels on each axis (images 0 and 3 in figure 1), with termination occurring within 5-10 steps. However, the actual convergence depends on the termination criterion. In this case, the termination criterion checks whether the Euclidean distance between the updated and previous coordinates is less than 0.5 pixels.

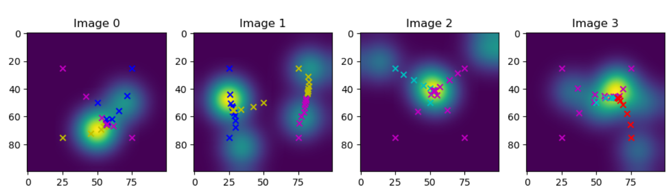

Figure 1. Mean-shift mode seeking on different responses.

B. Mean-shift tracking

Table I presents the failures per sequence, with an attempt to balance between good, mid, and bad runs.

The success of good runs is attributed to the object's distinctiveness from the background and sufficiently slow movement. For instance, the face sequence (Figure 2, left) illustrates that partial occlusions pose minimal challenges under slow movement conditions.

Mid runs encounter challenges due to occlusions, fast movements, and changes in object size, as observed in the zoom-in woman sequence (Figure 2, middle) and the bolt sequence. Consequently, occasional failures occur.

Bad runs are characterized by either rapid movement, as seen in the juice sequence (Figure 2, right), or a lack of distinction between the tracking object and the background, evident in the singer sequence with light glare or a moving background.

In summary, successful tracking requires constant or slow movement, a distinct tracking object, and preferably minimal occlusions.

Sequence	Fails
Cup	0
Face	0
Bolt	3
Woman	4
Juice	9
Singer	7

Table I
Fails per sequence

Figure 2. Some snapshots of the tracker (left slow occlusion, middle fast zoom in, right fast camera movement).

To achieve these results, assigning full weight ($\alpha = 1$) to the newly obtained histogram during tracking updates was crucial. Any weight given to the $t - 1$ histogram resulted in poor performance, particularly in mid-level sequences like bolt and woman, with over 20-30 failures, rendering tracking in bad sequences futile.

Adjusting the number of histogram buckets effectively controlled sensitivity to noise and lighting variations, with 16 buckets yielding optimal performance. Additionally, using 16 buckets reduced computational overhead compared to higher counts.

Kernel size played a pivotal role akin to histogram buckets, requiring precise tuning. A kernel size of 31 by 31 pixels yielded the best results, with deviations leading to a higher likelihood of tracking errors.

The termination criterion had minimal impact, as the mean-shift algorithm typically converged rapidly, with stricter criteria resulting in negligible changes in the final position.

Regarding computational efficiency, the tracker demonstrated high performance, achieving speeds of 500-800 frames per second depending on the sequence.

III. Conclusion

In conclusion, we've explored the mean-shift algorithm's parameters and its application in object tracking. Through experimentation, we've identified key factors influencing tracking performance, such as kernel type and size, histogram weighting, and termination criteria. These insights contribute to more effective and robust object tracking systems.